GÖTEBORG UNIVERSITY
FACULTY OF SCIENCE

RELATIONSHIPS BETWEEN NOMENCLATURE, PHYLOGENETICS AND SYSTEMATICS

YANN BERTRAND
Department of Zoology, Systematics and Biodiversity
Göteborg University
Sweden

DISSERTATION
Göteborg, 2008

ISBN 978-91-628-7544-2

Academic doctoral thesis in systematics and biodiversity that will, on the decision of the faculty of natural sciences, be publicly defended on Friday 26 September, 2008, at 10.00, in föreläsningssalen, Department of Zoology, Medicinaregatan 18, Göteborg. Faculty opponent is Professor Mieczyslaw Wolsan, Polish Academy of Sciences, Warszawa.

BERTRAND, Y.[1,2] Relationships Between Nomenclature, Phylogenetics and Systematics. ISBN 978-91-628-7544-2

[1]*School of Life Sciences, Department of Biology, Södertörn University College, SE-141 89 Huddinge, Sweden; E-mail: yann.bertrand@sh.se*

[2]*Department of Zoology, Systematics and Biodiversity, Göteborg University, P.O. Box 463, SE-405 30 Göteborg, Sweden.*

Abstract: Systematists have become increasingly aware of the limits imposed by the current system of nomenclature for accurately representing evolutionary relationships and managing efficiently names associated with clades. In reaction, a new system of nomenclature, the PhyloCode is being developed that fully recognizes the historical nature of taxonomy and the importance of the cladistics revolution. As a consequence, questions emerge about the new historical entities of systematics, questions that can be apprehended through the lens of epistemology, philosophy of language and metaphysics.

What is the ontological nature of entities that lack any other essential features besides spatiotemporal properties? How to depart from the fixed realm of immutable and transcendental essence into a worldview wherein all biological entities are characterized by their temporality and materiality? What are the consequences of nomenclatural decisions on other sectors of biology? With the ever growing sequencing capacity and tree reconstructing abilities, our conceptualization of phylogenetic relationships is changing at an unprecedented pace. Then it begs the question, what prevents communication break down when the references of clades' names are changing almost on a daily basis. These are some of the fundamental issues I am tackling in the present work.

Addressing the ontological issue, I argue that species and clades are best perceived as mereological sums of individuals, which means that each biological individual is the unique individual composed of all its less inclusive individuals and nothing more.

I propose to separate the meanings of "clade" and "monophyletic group". I suggest to use "monophyletic" for an epithet referring to a defining property of a set (a natural kind) and "clade" for a noun which corresponds to a historical entity (an individual) resulting from evolutionary process. I present the idea that a phyloname is not attached to a single clade but to a natural kind containing as members the clades that would be selected in counterfactual phylogenies. The defining properties of this natural kind are provided by the phylogenetic definition.

Finally I stress that taxonomists are also driven by the will to narrate the same sort of history, when they adjust the reference of names in light of new phylogenetic data, which leads me to submit that taxa can also be perceived as narratives.

Keywords: PhyloCode, philosophy, systematics, individuality, natural kind, possible worlds, causal theory of reference

LIST OF ORIGINAL PUBLICATIONS AND MANUSCRIPTS

This thesis is based on the following papers:

ARTICLE 1: BERTRAND, Y., PLEIJEL, F., & G.R., ROUSE, 2006. Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks. *Systematics and Biodiversity*, 4: 149-159.

ARTICLE 2: BERTRAND, Y., & M., HÄRLIN, 2006. Stability and universality in the application of taxon names in phylogenetic nomenclature. *Systematic Biology*, 55: 848-858.

ARTICLE 3: BERTRAND, Y. Species individuality and integration. Manuscript.

ARTICLE 4: BERTRAND, Y. 2008. Contrasting the general with the particular in phylogenetics - a proposal to keep the meanings of mono/paraphyletic and clade/grade separated, *Taxon*, 57: 705-708.

ARTICLE 5: BERTRAND, Y., & M., HÄRLIN, 2008 .Phylogenetic hypotheses, taxonomic sameness and the reference of taxon names. *Zoologica Scripta*, 37: 337-347.

ARTICLE 6: BERTRAND, Y., & M., HÄRLIN. Historicism and essentialism in phylogenetic biology. Manuscript.

CONTENTS

"…one significant point is that philosophers habitually seek only to erect internally consistent schemes. They have little to gain by attempting to discover how the living world really is organized, or how biologists in fact conceive of it. Hence although there has been much philosophizing *about* science, there has been little that truly deserves to be called the philosophy of science. The usual notions of the logicians are adequate enough that these can be applied to scientific classifications without breaking down. Something special was needed if logic were to cope with biology as it really is. This situation is unfortunate, for not only might philosophers have much to contribute, but they might learn something as well." (Ghiselin, 1974).

INTRODUCTION

How far does evolutionary thinking extend in biology? The question seems terribly naïve at first glance, since the appropriate quotation comes readily to mind: “Nothing in Biology Makes Sense Except in the Light of Evolution" (Dobzhansky, 1964). No doubt that the publication of *The Origin of Species* triggered a paradigm shift whose influence transcends disciplines’ boundaries, as Ghiselin (1969: p. 1) asserts emphatically: “In 1859 there began what ultimately may prove to be the greatest revolution in the history of thought. *The Origin of Species*, published in November of that year, affected an immediate and cataclysmic shift in outlook, casting into doubt ideas that had seemed basic to man’s conception of the entire universe.” During the 1940’s *The Modern Synthesis* extended the influence of the evolutionary theory from the mere origin of species and enabled to unite all biological disciplines. Dobzhansky’s (1964) manifesto is often perceived as the triumph of the evolutionary theory, which became the central principle in biology, or so goes the usual story. Yet, a moment’s reflection reveals such perspective not to be fully justified. If we cast a less partisan eye on history, we observe that despite its explanatory power, biologists were not long to embrace the molecular revolution and soon evolution lost its role as an integrative theory. Molecules, gene and DNA are today conceptual pillars in biology.

During the 1960’s cladistics and then later in the 1990’s the refinement of computing methods brought a renewed interest in the theory of evolution especially in systematics. As a consequence systematics is now considered as a mature science in the popperian sense, meaning it is a science able to generate and test phylogenetic hypotheses, but also to arrange these hypotheses into a system. As emphasized by

Griffiths (1976), modern systematics deals no more with classification, the arrangement of objects into classes on the basis of properties of the objects being classified. Instead, in its modern form, systematics handles a vast array of concepts and methods that purport to arrange the identified elementary objects into larger wholes. Genealogical relationships form the cement that enables the constituent objects to bear the relation of connected parts to the whole. The goal of modern systematics is to group organisms into taxa, which may or may not be monophyletic, and to rank those taxa into a hierarchy of taxonomic categories. Under the traditional rank based system, naming hinges on ranking, which determines the ending of the name, and on selecting a type, which provides the stem for the name. Furthermore the role of the type is to anchor a name into biological reality. At the end of the process every taxonomic object is endowed with a name, a rank and a type.

The introduction of cladistic philosophy instigated a change of perspective in the way systematists view taxa, at least at the levels above species, which hereafter correspond to monophyletic groups of species. However difficulties arise for the traditional system when the Tree of Life is perceived as a fundamentally continuous historical entity, which ends with the extant organisms and begins with their last common ancestor. Recognizing this continuity forces us to admit that ranks do no correspond to any natural feature and as a corollary any division of this continuum is *a priori* as valid as any other. Ranking does indeed rely on subjective judgments from the taxonomist who takes simultaneously into account different kinds of information such as the size of the taxon, the morphological gap to other taxa or the ranks of neighboring taxa. Because it generates endless debates on the appropriate level for a taxon in the hierarchy, ranking is a cause of nomenclatural instability. It is thus incumbent on the traditional system to

adopt a fully objective method for ranking, but the proposed solutions (see for instance Hennig, 1966; Sibley & Ahlquist, 1990; Avise & Johns, 1999) fail to conciliate objectivity with communication ease. An additional difficulty in this system pertains to the link between names and ranks, in which rank changes might induce cascading series of names changes (Hibbett & Donoghue, 1998).

In sum, systematists have become increasingly aware in recent years of the limits imposed by the current system for accurately representing evolutionary relationships and managing efficiently names associated with clades. In reaction, a new system of nomenclature, the PhyloCode (Cantino & de Queiroz, 2007) is being developed. Unlike the traditional system, the PhyloCode does not enforce any ranking requirement and concerns itself solely with the names of clades, leaving the thorny issue of species naming to traditional systems (Dayrat *et al*., 2007). This new system intends to provide clades both with unambiguous names and precise temporal boundaries. Under this new approach, systematists can freely select and name any part in a phylogeny, depending on their communication requirements and this without further concerns about its appropriate rank. The resulting classification is constructed out of clades included into other clades and thus reflects the shape of the Tree of Life. With the PhyloCode, the scientific community has the opportunity to extend tree thinking into nomenclature and by doing so to acknowledge the fully historical nature of systematics.

The main outcome of the Darwinian revolution, the importance of the mechanism of natural selection notwithstanding, was to introduce the concept of time at the heart of biological reflection. The prominent influence of *The Modern Synthesis* during the 1950s and the widespread acceptance of phylogenetic trees, that followed the cladistic

revolution, are all signs of the historical turn in biology. Interestingly the same trend has been observed in other conceptually remote disciplines during the second half of the 20th century. Most notably for our purpose, interest in the temporal dimension rose during the same era in epistemology, philosophy of language and ontology. In epistemology, Kuhn (1962) and Feyerabend (1975) appealed to history and sociology to show that the dynamics of science proceeds from factors other than purely cognitive. In philosophy of language, Putnam (1975) and Kripke (1980) theorized the persistence of communication despite the impermanence of the sense associated with words. Finally in ontology, Ghiselin, (1966) and Hull (1976) challenged the reigning view inherited from classical logic that entities in systematics are best thought of as natural classes defined by eternal essences. They proposed that they rather should be apprehended as individuals, that is, as historical objects devoid of necessary and sufficient properties. These three disciplines might provide us with the conceptual tools to prolong the Darwinian revolution in systematics and beyond in biology.

What is the ontological nature of entities that lack any other essential features besides spatiotemporal properties? How to depart from the fixed realm of immutable and transcendental essence into a worldview wherein all biological entities are characterized by their temporality and materiality? What are the consequences of nomenclatural decisions on other sectors of biology? With the ever growing sequencing capacity and tree reconstructing abilities, our conceptualization of phylogenetic relationships is changing at an unprecedented pace. Then it begs the question, what prevents communication break down when the references of clades' names are changing almost on a daily basis. These are some of the fundamental issues I am tackling in the present work.

During my thesis, I was driven by the excitement of exploring philosophical realms, very distant from what I had been acquainted with during my biological training. Yet, I was struck to learn that philosophers are interested in the very same issues I intended to address, which prompted my attempt to import some philosophical tools into systematics. As a result, it allowed me to deal with the above problems but also it poses new questions that might in turn challenge philosophers.

Although, Dobzhansky was premature in believing that the Gestalt shift had already occurred, he anticipated a direction toward which biology is currently heading. Consequently, my overall project is to foster the expansion of evolutionary thinking toward a biological frame in which every single object, from the units considered in conservation biology to the concept used by metaphysicians, would be historical in nature. This core ambition explains the breadth of subjects encountered in the present thesis. Measures of biological diversity, history of classification and phylogenetic nomenclature are some of the important concepts I have been manipulating. Furthermore, this work has the ambition to contribute in the design of coherent framework in systematics that shed light on the ontological nature of its entities and understanding on how these entities are causally related. Therefore this thesis should also be received as a work in metaphysics of biology.

Before turning to my own contribution, I wish to introduce the main theoretical concepts and tools I have appealed to. Firstly, an explanation of the PhyloCode's foundations is required in order to understand some problems resulting from its application. Secondly, I will present the two fundamental categories used in ontological claims, namely individuals and natural kinds. Thirdly, I will outline the causal theory of

reference, cornerstone of the contemporary philosophy of language. Directly related to the previous theory, the theory of possible worlds will be the focus of the last part. The possible worlds theory stems from the development of modal logic (i.e. the logic of necessity, possibility and contingency). Although this theory proved to have a high heuristic value, it is not unified yet and still encompasses number of dissenting views. Due to its controversial aspects I will not paint a complete representation of the theory and will only provide a short introduction to some of its most promising features to analyze modality issues in systematics.

THEORETICAL FRAMEWORK

THE PHYLOCODE

Most of the current draft's key features were already present in de Queiroz & Gauthier's (1990, 1992, 1994) introductory articles. As for any system of nomenclature the prime function of the PhyloCode is to provide its users with directions on which entities should be recognized in this system and once they are recognized on how to name them. At this stage, the PhyloCode concerns only the naming of clades, defined as monophyletic groups of species. Although it does not preclude the naming of para- and polyphyletic groups, the PhyloCode discourages it and their names would not be managed by this system. In the PhyloCode the choice of a phylogenetic reconstruction precedes naming. Admittedly, the systematist does not intend to name every single clade in the phylogeny, but focuses on the groups that are the most relevant for his purpose. Attention could be drawn on a particular clade because of its importance in past debates, the presence of some significant characters or because of its special position in the tree.

The main feature introduced by the PhyloCode is the phylogenetic definition. Though it could be argued that the current rank based system is also using some sort of implicit definitions based on ranks and types, the PhyloCode departs from it by uniquely associating a name with an explicit phylogenetic definition. The phylogenetic hypothesis that provides a context for coining the initial phylogenetic is called the reference phylogeny. When formulating the phylogenetic definitions, the PhyloCode uses specimens and characters present in the reference phylogeny that have been dubbed specifiers. They function analogously to types in the rank based system in anchoring a name into biological reality and thus providing reference points for inferring the extensions of names. Phylogenetic definitions purport to identify the referent of a name, in any phylogeny where the definition applies. In theory this identification is meant to be unambiguous even in cases when the hypotheses depart radically from the reference phylogeny.

The PhyloCode proposes a repertoire of phylogenetic definitions that enables to name any portion of a tree. It can be narrowed down to three main definitional types, which might in turn be combined with each others. Using the specimens A and B and the character state M as specifiers, examples of referent clades associated with each type of definition are for a node-based definition, “the clade stemming from the last common ancestor of A and B”, for a branch-based definition, “the clade consisting of A and all organisms or species that share a more recent common ancestor with A than with B” and for an apomorphy-based definition, “the clade originating with the first organism or species to possess apomorphy M as inherited by A". As a consequence a name is attached to a definition and in ordinary practice the name’s referent is identified automatically

without requiring any taxonomic judgment. However, an extraordinary situation corresponds to the acceptance of hypotheses, wherein circumscriptions of taxa are profoundly modified in comparison with the reference phylogeny. In such cases, the systematist might sever the link between a name and a definition, and propose a redefinition that aims to capture the conceptualization in term of the taxon content initially intended by the author of the definition. The systematist has two options for intervening on the wording of a previously published definition. The first possibility is *an unrestricted emendation* that is a modification, which contrary to *a restricted emendation*, do not require approval by the elected Committee for Phylogenetic Nomenclature (CPN). Unrestricted emendations may modify the specifiers or the qualifying clauses, but must retain the same definitional type and the same clade category (crown clade or total clade) if specified. The second possibility may involve changes in specifiers, definition type, clade category and/or definitional type. One of the central principles in the PhyloCode is that it does not infringe upon the freedom of taxonomist with respect to inferring the composition of taxa. As a consequence, it is silent on the requirements for making emendations, stating merely that they should be derived from an interest of stability in terms of taxon composition and/or diagnostic characters (Article 15.1.1). This state of affair seems to me quite unfortunate especially concerning unrestricted emendations. Indeed without precise guidance, systematists might revert to the old habit of changing taxon definitions to fit their favorite circumscription. Due to the shape of the tree, in many cases the exact taxa composition cannot be maintained and several solution do equally approximate it. Thus, as demonstrated in articles 2 and 3,

conflict could arise about the best emendation, bringing havoc once again when we thought that nomenclatural stability had been achieved.

INDIVIDUALS AND NATURAL KINDS

Since its origin, metaphysics has been concerned with different ways of aggregating objects of the world in order to give us ways of thinking about many things as one thing. Individual and natural kind form the fundamental ontological pair, each of them captures a different part of reality. In that respect they should be perceived as a pair of contrasted terms each of which depends on the other for its meaning. Although, the words used for describing this opposition have varied, the associated concepts have remained remarkably stable over time, and represent modern developments of the Aristotelian theory of universals.

Natural kinds are defined by essences, which correspond to sets of necessary and sufficient properties. According to the classical approach, such essences are universal, immutable and eternal - they form the very matrix of reality. Consequently, a natural kind exists transcendentally, i.e. whether there are concrete objects instantiating it or not. Among mildly ill-informed logicians with respect to biological things, biological species are still considered to be paradigmatic cases of natural kinds (see Kriple, 1980). In biology, before Ghiselin (1966) formalized an alternative, the approach that treated species as natural kinds was the dominant view. To have the knowledge of the species tiger was to discover its essence. Candidates for essential properties were morphological traits, biochemical and genetic features (Bernier, 1984; Caplan, 1981; Kitts & Kitts, 1979 and Kitcher, 1984). *Hélas*, once all the implications of the theory of evolution were taken

into account, the previous properties did not fit anymore the classical version of essentialism. Contrary to a widely held opinion, essentialism was a working hypothesis even after the acceptance of the evolutionary theory. As Sober (1980; 2000) demonstrates it, essentialism was able to accommodate itself with the gradual transformation of species in a similar manner that the transmutation of chemical elements did not undermine the periodical table. The fatal blow came from Galton's insight that species were not made up of homogeneously populations and that heredity was a central cause of variation: "…variability within one generation is explained by appeal to variability in the previous generation and to facts about the transmission of variability" (Sober, 1993: 174). For classical essentialism, variation was due to perturbations from an intrinsic natural state and to use Sober's famous terminology, it had to be *explained away* in order to reach the true underlying necessary properties. This attitude contrasted with the newly developed framework of population genetics wherein variability *per se* was the subject of investigation. Furthermore, evolutionary arguments against the sufficiency requirement started to pile up once it was understood that convergence and inheritance of ancestral characters cause identical features to be present among distantly related species (Ereshefsky, 2001). The version of essentialism, which relies on intrinsic properties of organisms, felt in demise as it was realized that it was ill suited for the ontology of systematics. Notice that it does not entail that essentialism has no place in biology, merely that species and clades do not conform to the classical picture of natural kinds.

In contrast with natural kinds, the individuality thesis is not beset by the difficulty of identifying essences. This thesis states that species and clades are individuals (Ghiselin, 1966), which means they are concrete particulars that lack instances and

defining properties. They are historical entities that form spatiotemporally integrated and cohesive wholes (Ghiselin, 1974). The individuality thesis is currently the dominant position, but some have argued that it could be complemented by treating species and clades also as natural kinds. The authors of this proposal have widened the notion of essence to capture the historical aspect of organisms' properties in the Darwinian era. The rational for reintroducing natural kinds is that they are kinds with explanatory and predictive values. According to LaPorte (2004: 19), many features of, for example, a polar bear can be explained and predicted by appealing to natural kinds it belongs to: "A lot is explained by an object's being a polar bear. That is a polar bear explains why it raises cubs as it does, or why it has extremely dense fur, or why it swims long distances through icy water in search of ice floes. Similarly, there are theoretically satisfying answers as to why polar bears on the whole raise cubs as they do, or have dense fur, or swim for miles through icy water. The polar bear kind is a useful one for providing significant explanations. It is a natural kind." The two main candidates for defining historical natural kinds are essence associated with shared ancestry and essence that posit homeostatic mechanism. The former have been defended by Griffiths (1994) and LaPorte (2004). LaPorte (2004: 15) argues that any talk about a species individual could also be interpreted as a talk about a natural kind, because the organisms that are part of the species individual also share a common property: "Although the species-individual is not a kind but rather an individual, there is a property, for any such individual, of being part of that individual". Both cladistic taxa based on purely genealogical considerations and the taxa of evolutionary taxonomists that account also for other type of information such as ecology and genetic differentiation count as perfectly cogent natural kinds for LaPorte.

Their naturalness consists in their explanatory values, but LaPorte stresses that naturalness comes in degree. The cladistic taxon is more natural when it comes to explain the genealogical relationships, whereas the evolutionary taxonomists' taxon is more natural when there is a tradeoff between genealogical and say ecological explanations.

The second category of natural kinds is defined by what Boyd has named *homeostatic properties*. Similarly to LaPorte's account, taxa (species and clades) are not defined by necessary and sufficient conditions, are not necessary spatiotemporally unrestricted, do not fall under universal exceptionless laws, and have their members united by historical causes rather than by shared properties. But, for Boyd (1999) the causal relations uniting the members of a species are homeostatic properties exemplified by gene exchanges between populations, reproductive isolation, effects of common selective factors, coadapted gene complexes plus other limitations on heritable variation, and the effects of ecological niches. Rieppel (2005) extended the concept to encompass clades, wherein the homeostatic mechanisms should be searched among developmental homeostasis, developmental constraints and ontogenic entrenchments. Once the thesis that taxa are homeostatic properties clusters is adopted, the question remains whether they are also natural kinds. According to Boyd, the naturalness of natural kinds consists in their aptness for induction and explanation. Therefore the identification of natural kinds amounts to the identification of projectible generalizations (in the sense of Goodman, 1973), which concerns the property of certain generalizations to be extended or projected from examined instances of the kind to unexamined ones. Since our ability to identify species and their members is central to our aptitude to obtain correct

explanation and predictions in biological sciences, Boyd concludes that taxa defined by HPC are natural kinds.

In the articles 4, 5 and 6 we build upon these accounts of historical essences but explore an alternative version of natural kinds more compatible with the individuality of species and clades.

THE CAUSAL THEORY OF REFERENCE

Associated with the works of Saul Kripke and Hilary Putnam, the causal theory of reference spawned a shift in philosophy of language in the early 1970s. This new theory of reference is motivated by the rejection of orthodox philosophy of semantics, held by Frege and to some extent by Russell, that certain expressions, in particular proper names, are non-connotative expressions. I will focus hereafter on Kripke's account as it is exposed in his most influential work: *Naming and Necessity* (1980). The version of the theory of reference, attacked by Kripke, is the descriptive theory of reference. It is generally admitted in semantics that the way we determine the reference of an expression is by finding the reference of each of the terms that compose it. According to the descriptive theory every singular term receives, in addition to a denotation, or the object that the term refers to, a connotation, or a sense that provides the description of the referent. The sense of the term, thus, contains the features for identifying uniquely the referent of the term.

In what follows, I will mainly discuss the issue of proper names used to pick out individual particulars. Kripke argued that the association of proper name with a single property might fail to provide necessary and sufficient conditions to identify a unique

referent. Rather, he proposed that the descriptive theory would be better off attributing to the referent a disjunctive cluster of properties. In turn, this cluster would be necessary and sufficient to identify the referent of the name as summarized by Kripke (1980: 71): "If most, or a weighted most, of the φ's [cluster of properties] are satisfied by one unique object γ, then γ is the referent of '*X*' [the proper name]. If the vote yields no unique object, '*X*' does not refer." For Kripke, this forces us to accept that this cluster of properties is essential to the entity that answers to the description. For instance, we identify the referent of 'Aristotle' by way of a description that identifies him uniquely (Greek philosopher, pupil of Plato, teacher of Alexander the Great) so it seems as an inevitable consequence that this description is based on Aristotle's essential properties, those characteristics that make him Aristotle and nobody else. Kripke's arguments against this version of a theory of meaning can be decomposed into modal and epistemological arguments (Salmon, 2005).

The modal argument appeals to counterfactual situations and purports to show the fallacy of "the statement, If *X* exists, then *X* has most of the φ's' expresses a necessary truth" (Kripke, 1980: 71). Suppose that the following description "the man who won the American presidential election in 1968" gives the essential property of Nixon and that the name "Nixon" simply means "the person, whoever he or she may be, having these properties", then according to the descriptive theory of reference, the sentence "Nixon is the man who won the American presidential election in 1968" is analytic. An analytic statement is, for Kripke, true in all counterfactual situations by virtue of its meaning. Consequently Kripke stipulates (p. 39): "Then something which is analytically true will be both necessary and *a priori*". But, surely the argument continues, it is not necessary

that someone is Nixon if and only if that person is the man who won the American presidential election in 1968. In the first place it might have come to pass that someone else (e.g., Humphrey) won the elections in 1968. Hence the above sentence expresses a contingent truth, thus it not necessary for Nixon to have the property that we have attributed to him.

The epistemological argument attacks the alleged *a prioricity* of descriptive statements summarized by "the statement 'If *X* exists, then *X* has most of the φ's' is know *a priori* by the speaker" (Kripke, 1980: 71). As we know from the previous section, "if anyone is the winner of the American presidential election in 1968, then he is Nixon" is assumed analytic by the orthodox theory, which means that this sentence conveys information that is knowable *a priori*, knowable solely by reflection on the concept and without recourse to sensory experience. Yet, Kripke argued that such sentence might prove to be wrong that our memories are tricking us into forgetting that the winner of the 1968 election was in fact Humphrey. It follows that the information conveyed by the sentence is in fact knowable *a posteriory* by recourse to sensory information, for instance by digging into the archives to check who was elected president in 1968. Therefore this second sentence is not analytic either.

Together the modal and epistemological arguments showed that there are serious flaws in the claim that the referent of a proper name has any of the properties contained in its description commonly associated with it, let alone necessarily. Kripke proposed a better picture dubbed *Causal theory of reference*, which says that a name denotes a referent but does not connote any properties. In this scheme, the manner in which names are bestowed can be likened to christening by ostension, by the act of pointing or by

using a definite description. However this initial description plays only a role in fixing the reference of a name, it doesn't convey any necessary and sufficient properties and can even be wrong. Once established, the reference of a name is spread to the rest of the linguistic community (Kripke, 1980: 96): "When a name is 'passed from link to link', the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it. If I hear the name 'Napoleon' and decide it would be a nice name for my pet aardvark, I do not satisfy this condition". In this perspective, no matter the modal considerations we can imagine, no matter how our knowledge of the referent's properties might be modified, the name would always pick out the same person. According to Kripke we can surrender the defining properties entailed by the descriptive theory of reference because proper names are *rigid designator*, they are names with the ability to refer to the same bearer in all possible worlds. The concept of rigid designator is intensively used most of the articles, but is central in the articles 5 and 6.

THE POSSIBLE WORLDS THEORY

There are several versions of possible worlds theories. The most influential formalizations are due to Kripke (1963; 1980) and Lewis (1968; 1986), although there are points of contact between them, the details pertaining to each theory and the motivations that lie behind each one differ. In *Naming and Necessity*, Kripke resorted to a version of the theory that takes possible worlds to be heuristic devices in the study of modality. According to it, possible worlds are stipulated not discovered. Each possible world corresponds to a counterfactual situation, to the world as we know it. Kripke explains

their construction (p. 44): “A possible world is given by the description we associate with it. What do we mean when we say ‘In some other possible world I would not have given this lecture today?’ We just imagine the situation where I didn’t decide to give this lecture or decided to give it on some other day. Of course, we don’t imagine everything that is true or false, but only those things relevant to my giving the lecture; but in theory, everything needs to be decided to make a total description of the world”.

According to the causal theory of reference, the truth-value of a sentence is a function of the truth-value of its parts. The only feature relevant to the truth-value of a single part is its referent. To assert something, amounts to designate a referent and to ascribe it a property. Hence in the proposition “Dunkerque is a harbor”, the proper name refers to a geographical object, the city of Dunkerque, whereas the common noun expresses the property of being a harbor. Thus the proposition is true, if and only if the referent of the name “Dunkerque” possesses effectively the property of being a harbor. In the possible worlds language, the referent of “Dunkerque” is an actualized possibility, because the actual city is present in the world you and me inhabit. But what about a modal proposition such as “if Dunkerque was totally destroyed during the war, it would have been necessarily rebuilt inland”? In this context, “Dunkerque” has no actual referent, because the city was not totally destroyed during the war. This begs the question how we should decide about the truth of this second proposition. This is where the mechanic of possible worlds springs into action. The solution consists in stipulating some possible worlds as close as possible from the actual world in every respect except that Dunkerque was totally destroyed during the war. Then each of the counterfactual worlds is examined in order to see if it includes the city of Dunkerque rebuilt inland. If in every

single possible world Dunkerque has been rebuilt inland, the reconstruction would have taken place necessarily inland, which entails that the proposition is true. If in some worlds the reconstruction is inland, whereas in some others it is localized on the seashore, the proposition is wrong. But the proposition that substitutes “possibly” for “necessarily” is true. For Kripke a proposition that is true in every possible world where it applies is said to be metaphysically necessary.

In the context of my work, the theory of possible worlds was used to investigate the necessary features of systematics’ objects. The procedure consisted in exporting each possible phylogeny in a world of its own and then in comparing the different worlds in order to discover which features were always present, i.e. which marked necessary conditions.

However there is a second potential benefit to use the possible worlds framework in phylogenetics, which I wish to explore further in the future. This idea, introduced by Lewis (1986) (and based on the work of R. Hilpinen, 1976) is centered around the premise that to come up with a good scientific theory is tantamount to find an explanation that is close to the truth. This closeness to the truth (“truthlikeness” or “verisimilitude”) is formalized in term of possible worlds closeness. Thus, Lewis writes (1986: 24): “A theory is close to the truth to the extent that our world resembles some world where that theory is exactly true. A true theory is closest to the truth, because our world is the world where the theory is true. As for false theories, the ones that can come true in ways that involve little dissimilarity to the world as it really is are thereby closer to the truth”. Lewis example, the approximate gas laws, helps to clarify this inferential model. When the laws are used without including correction terms the worlds where it

perfectly describes gas behavior (the worlds where the laws are true) are quite close from our world. But when we include some corrections "we get a theory that holds in some worlds that imitate ours still better, so the improved theory is still closer to the truth". Consequently, this means that science proposes theories that hold true in some counterfactual worlds, it also means that we don't know *a priori*, i.e. before testing them empirically, whether our world is the world where those theories are true or how close are these counterfactual worlds from our own. It emerges from this analysis that empirical science can be understood as the process of leafing through the atlas of *possibilia*, selecting one and confronting this possible world to the observation in order to infer (deductively, inductively, abductively or with an inferential mix of the previous modes) whether it is the actual world or how close it is from the actual world. I call *actualization* the process of selecting one *possibilia* as the world we actually live in. Of course, as Popper has thought us, this selection is only provisional since we cannot be certain that different theories, i.e. better *possibilia*, will not provide better accounts of the observation data. Scientific theories are entangled and cohesive explanatory wholes emerge from their interactions. Hence a modification of a given part might exert changes in verisimilitude in a wide range of other theories. Thus, some particular theory, we believed in when embedded in a given theoretic context, might show a drastic change in verisimilitude when reconsidered in a different theoretical environment.

In the possible worlds context, the aim of phylogenetics is not to reconstruct a tree, but to select a tree (or several equally "good" trees), i.e. in my terminology to actualize a tree, from a set of all possible trees. The selection procedure consists in confronting the phylogeny found in each possible world with the theoretical entities that

constitute the matrix of primary homologies. Therefore, actualization in phylogenetics is a form of explanation that purports to infer the causes of the observed primary homologies.

The four theories and theoretical frameworks introduced above have been deeply influential in my work. With hindsight they have proved very efficient in disentangling metaphysical issues in systematics. I have always attempted to discover how I could weave interactions between them. However, I have also been confronted to situations when these philosophical tools do not apply readily to systematics. Instead of discarding them, I choose to interpret them in a philosophically unorthodox way. This is the reason why some of the following articles are not confined to systematics and should also be perceived as philosophical contributions.

SCIENTIFIC ARTICLES

The first article of this compendium addresses some problems that stem from the inappropriate use of the Linnaean classification. The next articulates some objections against recent recommendations on increased stability in the PhyloCode. In the reminding articles I work inside the framework of phylogenetic nomenclature without accepting all the conventions imposed by the PhyloCode, most notably and for some obvious reasons in the paper which bears on the species issue.

ARTICLE 1: Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks, **Y. Bertrand,** F. Pleijel & G.R. Rouse, 2006, *Systematics and Biodiversity*, 4: 149-159.

Contrary to the general feeling that could arise upon reading this article, I want to state from the outset that it does not advocate rejecting the Linnaean system in favor of the PhyloCode. It conveys a rather different message, that the Linnaean groups cannot fulfill some tasks they were not designed for in the first place. In particular, it castigates the method known as *taxonomic surrogacy*, which is based on the idea that species numbers need not be assessed directly, but can be estimated indirectly by counting higher ranked taxa, such as genera, families or orders. Taxonomic surrogacy has been advocated by a number of authors (e.g., Andersen, 1995; Balmford *et al.*, 1996a; Balmford *et al.*, 1996b; Balmford *et al.*, 2000; Gaston & Williams, 1993; Lee, 1997; Warwick, 1988; Williams & Gaston, 1994) in several quite unrelated disciplines such as palaeontology, for identifying the magnitude of extinction events, in community perturbation, for detecting pollution effects and in conservation, for predicting patterns of biodiversity hotspots.

The taxonomic surrogacy approach is based on the premise that there are predictable relationships between species and higher taxonomic ranks and that the higher taxon counts allows predicting species richness with a reasonable amount of precision. Suppose that we know the species richness for a number of sites. Based on existing classifications we can, for each site, compare species richness with, say, family richness and then calculate the regression line giving family richness as a function of species richness. This relation is conjectured to hold constant for other taxa and other sites, where in turn species richness is estimated from family richness.

In our article we demonstrate that there are no empirical, nor theoretical, reasons for the above relation between species and higher taxa to remain constant enough to be of practical value. We trace the origin of the method to an erroneous belief in some sort of

equivalence between taxa of the same rank, which misrepresents the underlying tree and the historical nature of taxa. Thus, one genus in one part of the tree is not equivalent to another genus in another part of the tree, no more (and no less) than the genus *Rattus* is equivalent to the phylum Arthropoda; these names merely make reference to different historical events. The only information present in two taxon names of the same rank is that they are non-nested.

Moreover this method fails to recognize that knowledge of the classification is not an incremental development toward the most accurate classificatory scheme of the natural world. Quite the opposite in fact, the classification is characterized by a large historical inertia, keeping in its taxa the signs of past theories that researchers were committed to. Examples of taxa recognized in the early era of systematics that are still in use today abound. For instance, the works of de Candolle, Jussieu, Bentham and Hooker still influence contemporary botany to a high degree. Augustin-Pyramus de Candolle (1778-1841) believed that variation in Nature was discontinuous and therefore taxa should be recognized as discrete and real. He emphasized that the numbers of individuals in species, species in genera, or genera in families, did not affect the ranks of those taxa and suggested ranking should be based on an equivalent value of morphological variation. In other worlds, the included variation in each family should be equivalent, albeit higher than the amount of variation that characterizes genera (Stevens, 1994). In contrast, Bentham and Hooker in their *Genera plantarum*, were actively dismembering large groups and lumping small groups, deemed respectively too large to be conveniently handled and too small to be informative. Another influential botanist of the 19th century is Antoine-Laurent de Jussieu (1748-1836), who conceptualized relations between taxa as a

continuous and linear *scala* of increased perfection. He argued that due to the continuity of Nature, systematists could manipulate the size of groups at will and any delimitation was as justified as any other, which entailed that the size of taxa was based primarily on subjective convenience criteria.

This short review illustrates how taxonomic philosophy has varied over time, but it could still be argued that a classification progressing toward the true representation of relationships in nature might emerge from the accumulation of observation data. However Stevens (1994: 487) casts serious doubts on this view when he explains that "[A]lthough new data have been added to classifications over the years, given that there is neither any clear way of evaluating data nor or representing the data on which a particular classification was based, relating classification to observations is no simple task. Knowledge of character variation is not cumulative; knowledge can also be lost."

Although this paper is coated in a practical disguise, it is first and foremost an attempt to spell out the position that systematics is guided by certain personal orientations and contextual developments in the field, which traces are trapped in the classification and passed from one generation of scientists to the next.

ARTICLE 2: Stability and universality in the application of taxon names in phylogenetic nomenclature, **Y. Bertrand** & M. Härlin, 2006, *Systematic Biology*, 55: 848-858.

One of the main issues opposing proponents of the phylogenetic system and defenders of the rank based system pertains to the problem of content stability, i.e. stability in the extension of names. The claim that the rank based system is more stable than the phylogenetic system is a recurrent critic (Moore, 1998; Benton, 2000; Nixon and

Carpenter, 2000; Dyke, 2002; Carpenter, 2003; Kojima, 2003; Moore, 2003; Nixon *et al.*, 2003) which drove into the PhyloCode the addition of mechanisms that should promote content stability: crown-clades, qualifying clauses and the inflation of specifiers.

In this article we argue that the recent overemphasis on stability in the PhyloCode contradicts its original precepts and hampers our abilities to unambiguously identify the taxa referred to by the phylonames. We focus on the inflation of specifiers which is the practice of using multiple specifiers in phylogenetic definitions. It is based on the assumption that the more specifiers that are included in a definition the more stable in terms of taxon content the reference of the name tends to be, i.e. the name will necessarily refer to all its specifiers. By using names defined with large number of specifiers, we show that in an overwhelming number of published phylogenies on these taxa, their names could not be applied unambiguously, that is, their exact circumscriptions were unclear. The reason is that phylogeneticists do not necessarily use the same taxa in their analyses as the one appearing as specifiers. Although, it may be argued (Cantino & Olmstead, 2008) that it is not necessary for all specifiers to be present in order for a name to be applied, we believe such an argument to severely impoverish the clarity of phylogenetic nomenclature. That is, too many applications are possible for each name in a single tree. This suggests that the door is wide open for ambiguity and thus threatening the very reason for moving to phylogenetic nomenclature in the first place. We conclude our article by recommending restraining the inflation of specifiers in phylogenetic definitions.

Our article was criticized by Cantino & Olmstead (2008), who defended the position that external knowledge of the phylogenetic tree can provide information

about the position of missing specifiers. For instance, despite that the tree in Kaufmann and Wink (1994) does not include pollen morphology in the data matrix, *Nepetoidea* (defined as "the most inclusive crown clade exhibiting exhibiting hexacolpate pollen synapomorphic with that in *Nepeta cataria*") can still be applied to a part of the tree based on the distribution of hexacopate pollen in Lamiaceae tabulated by Cantino and Sanders (1986). In this case, pollen distribution would be used as a surrogate for locating a character transformation on the tree. Since an optimized character was not part of the search for character congruence during the phylogenetic analysis, it is not logically equivalent to a synapomorphy, but this is a minor quibble. We do not dispute that a knowledgeable taxonomist is fluent in the literature pertaining to his group and has no difficulties to draw a cogent conjecture about the position of missing specifiers.

Instead, what our paper argues, is that the vast majority of users of taxonomic names, now and in the future, are non-specialists lacking precisely this taxonomic expertise, i.e. they are taxonomic and nomenclatural consumers (not producers). As we see it, a major gain of switching to phylogenetic nomenclature comes from the automatic feature of taxon re-identification. Such simplicity is necessary when loads of new molecular phylogenetic hypotheses are produced every day while, at the same time, trained taxonomists are cruelly lacking. Put bluntly, there just can't be a taxonomist behind every phylogenetist's shoulder. The crucial point is that we are not downplaying the role of taxonomy in modern systematics, rather the opposite! Taxonomy operates during the establishment of the initial phylogenetic definition and allows for periodic readjustments that are necessary to realign new phylogenetic insights with the needs of the scientific community. These adjustments take the form of modifications in

phylogenetic definitions (restricted and unrestricted emendations, see the article 15. in the version 4b of the PhyloCode) that are now allowed by the PhyloCode and defended already by Härlin (1998).

In conclusion, the taxonomic expertise that establishes the definition and the application of the definitions by the end users, are two different aspects of biological systematics that do not concern the same categories of researchers. Our point of view is that when coining their definitions taxonomists should take into consideration the users' needs and practices.

ARTICLE 3: Species individuality and integration, **Y. Bertrand**, manuscript.

Ghiselin's (1966) seminal work is not just a key text in philosophy, it shacked the very foundations of biological ontology inherited from classical thinkers. He proposed that species should be understood as individuals rather than natural kinds. He drew several conclusion from this view: "If species are individuals, then: 1) their names are proper, 2) there cannot be instances of them, 3) they do not have defining properties (intensions), 4) their constituent organisms are parts, not members"(Ghiselin, 1974: 536). Two decades of debate on the proper requirements for individuality allowed Mishler and Brandon (1987) to adduce four properties for individuals: spatial boundedness (having edges), temporal boundedness (having beginnings and ends), integration and cohesion. Integration and cohesiveness have later been questioned by Ghiselin (1997) who suggested that individually, entails merely spatiotemporal boundedness and temporally continuity.

Lee and Wolsan (2002) launched a line of criticisms when they argued that without integration as a necessary property of individuals, Ghiselin's redefinition leads to a paradox. They reasoned that if individuals are not necessarily integrated entities, then lower level individuals that are composing species need not be integrated either. That is, if the loss of integration between organisms in a species does not mean the end of the species then the loss of integration between the cells of an organism does not mean the death of the organism. And, since the lowest level individuals of a species are eternal fundamental particles, we are forced to agree that species are eternal as well. Therefore, if species are eternal, they cannot be spatiotemporally restricted, and consequently they cannot be individuals.

In this paper, I present an argument that counters Lee and Wolsan's proof and still preserves Ghiselin stricter definition of individuality. This argument explores the thesis that species (and clades) are best perceived as mereological sums of individuals (Brogaard, 2004), which means that each biological individual (in Ghiselin's restricted sense) is the unique individual composed of all its less inclusive individuals and nothing more, i.e. any part of it overlaps one or more of the individuals that are composing it (Lewis, 1986). These less inclusive individuals are, just to name a few, the population individuals, cellular individuals, genetic individuals. The species-as-mereological-sum thesis (MST) allows both to retain the stricter definition of individuality and to argue that integration is a necessary feature of species. However in my framework, integration does not occur synchronically between the organisms in the species, but diachronically between all the lineages that together make up the sum, i.e. behavior, organism, cell, chromosome, gene lineages, etc.

Furthermore, I discuss in this article the consequences of the mereological sum thesis when applied within the framework of the General Lineage Species Concept (GLSC) (de Queiroz, 1998, 1999, 2005, 2007). The GLSC aims to unify diverse contemporary views on the nature of species by suggesting that they all agree on the underlying species ontology, namely, that species are segments of metapopulation-level lineages, whereas they disagree on the properties used to recognize the lineages. Although de Queiroz's intuition brings some light in the species debate, it also implies that the current species definitions are all valid in the sense that they point to different evolutionary events. Hence the GLSC is committed to a form of pluralism with respect to the phyly (mono/para/polyphyly) of species taxa. In the MST framework, the cohesion of the sum results from the presence of a particular lineage that unites organisms into a whole. The causal lineage is always continuous, in contrast the organism lineage might present some discontinuity and lead to paraphyletic groups of organisms. The MST also emphasizes that since, ontologically speaking, species are not different from clades they should be allowed to form a hierarchy of more or less inclusive, possibly overlapping species taxa. Admittedly such pluralism cannot be embraced in the framework of the Linnaean system and I end the article with a plea for a whole different approach of species nomenclature.

ARTICLE 4: Contrasting the general with the particular in phylogenetics - a proposal to keep the meanings of mono/paraphyletic and clade/grade separated, **Y. Bertrand**, 2008, *Taxon*, 57: 705-708.

This short article moves on to explore the ontology of supra specific taxa in the framework of phylogenetic nomenclature. My analysis begins with the observation that clade and monophyletic group are often used as two interchangeable terms, yet from their uses one can notice that they sometimes single out distinct objects. I propose to separate their meanings in order to direct attention on two fundamental kinds of entities in systematics, which mark the difference between individuals and natural kinds. Along these lines I suggest to use “monophyletic” for an epithet referring to a defining property of a set (a natural kind) and “clade” for a noun which corresponds to a historical entity (an individual) resulting from evolutionary process.

There is a large agreement on the concept of monophyly which is defined as “[T]he condition that a taxon or other group of organisms contains the most recent common ancestor of the group and all of its descendants…” (Hickman *et al.*, 2006). However the meaning of clade is less consensual. I suggest it should apply to a single conceptualization of biological history, which implies that a clade is defined by its conjectured ancestor, its hypothesized topology and branch length. Modifying any of these features amounts to selecting a different clade.

According to this proposal, to hold that (A, B, C) is a monophyletic group is to predicate this group with the property “contains an ancestor and all the descendants of that ancestor”. Thus, monophyly is a relational property that hypothesizes a particular evolutionary relation between taxa inside a group. But, a hypothesis of a monophyletic group encompasses several hypotheses of evolutionary history between the taxa. That is, according to the account presented here, it contains several possible clades, each of which had originated with a different ancestor. To talk about the monophyletic group (A, B, C)

amounts to discourse about a set containing the following clades ((AB)C), (A(BC)), ((AC)B) which do not merely represent different topologies, but really different hypotheses of history. I conclude that monophyly could be understood as the defining property of a natural kind whose members are clades.

ARTICLE 5: Phylogenetic hypotheses, taxonomic sameness and the reference of taxon names, **Y. Bertrand** & M. Härlin, 2008, *Zoologica Scripta*, 37: 337-347.

The conceptual distinction developed above, between a clade and a monophyletic group paved the way for understanding how we are able to communicate with each other and this despite the fact that the names' references, namely the clades figuring in phylogenetic hypotheses, change frequently when trees are revised. This means that the extension of a name is modified every time the phylogenetic hypothesis wherein the name can be applied is altered. Then a scientist using the name at some time before the change is not saying the same thing as a scientist using it at a time after the change, even when they utter exactly the same sentence. The upshot is that if the entities, to which the name corresponds, are different it bears on the question whether we should keep the name constant. In the PhyloCode's framework we have two options. The first possibility keeps the definition constant and associate it with a different name. This is done in order to underline that we are dealing with two distinct entities belonging to two different phylogenetic contexts. The second possibility keeps the same name for different phylogenies, but emends the definition to capture a similar extension. Although this latest option is now possible according to the article 15 in the PhyloCode, I want to call attention to the fact that the identity of clades implies more than mere correspondence in

term of content. For two clades to be identical they must correspond to the same evolutionary history, that is, they need also to have the same internal relationships. Here enters Kripke and the causal theory of reference. Kripke argued that although the extension of a theoretical term might vary over time its meaning remains stable. The reason is that its meaning do not stem from its description, but is given by its referent. Moreover because theoretical terms such as the names of individuals and natural kinds are rigid designators, their associations with their referents are kept constant in time. As we have seen it these associations are established by ostension. This picture implies that scientists can fix the extension of kinds while knowing next to nothing about their properties and can learn about them after the kind was dubbed. It also implies that latter scientists might correct past workers' views, as explained by Kripke (1980: 121): "…so we might also find out that tigers had *none* of the properties by which we originally identified them. Perhaps *none* are quadrupedal, none tawny yellow, none carnivorous, and so on; all these properties turn out to be based on optical illusions or other errors…" Therefore the name is always fixed to the same entity but our conceptualization of the entities' properties varies along with scientific discoveries. For Kripke the name is *in fine* attached to an essence which causes the existence of the very entity. Kripke thought that species should be considered as natural kinds, but did not propose any clear essence for them. Other philosophers hold that the essence of a species is found in the species-specific microstructure of its members. Putnam (1975), for example, speculates that the essence of lemons is the unique genetic code of lemons.

In the case of names defined by phylogenetic nomenclature, the Kripke/Putnam thesis would imply that we fix a name to a real clade, i.e. part of the tree of life as it

effectively evolved. Then phylogenetics endeavors to match the real tree via successive ameliorations that refine our best phylogenetic hypotheses. In this paper we argue that this is a wrong picture of taxonomy: Even if the tree of life is real, it does not necessarily follow that taxonomy deals with real clades, only with hypotheses about clades. New phylogenetic hypotheses will be advanced, some will require extensive readjustments and some will be thrown aside, without any hints on how far off they are from the real ontological tree. Once it is recognized that systematics deals with hypotheses, we can investigate what is kept constant, what could serve as a stable reference across hypotheses. According to our analysis, it is blatantly clear that none of the components usually associated with a clade, namely an ancestor, a tree topology or taxon content, are necessarily constant across hypotheses.

So what is the referent of a phylogenetic name? Should we discard the causal theory of reference in systematics and search for a different explanation of the persistence of communication despite hypotheses fluctuation?

The solution we propose is inspired by the previous article and the differences between clades and monophyletic groups. We suggest that a phyloname is not attached to a single clade but to a natural kind containing as members the clades that would be selected in counterfactual phylogenies. The defining properties of this natural kind are constructed out of the specifiers, the type of definition and information such as qualifying clauses, all provided by the phylogenetic definition.

An example might clarify this new role for natural kinds in systematics. Consider *Cladia*, defined with a node-based definition as “the least inclusive clade containing specifiers A and B”, then the name *Cladia* is *fixed by* the two specifiers and the node-

based definition. With the additional terminal C, the possible hypotheses containing these specifiers are (AB), (A)(BC)), and (B)(AC). Therefore, A, B and the phylogenetic definition *fix* the name *Cladia* to the natural kind [(AB), (A(BC)), (B(AC)], whose defining properties are A and B performing as specifiers in a node-based definition. This class is then the referent of the name. Depending on the phylogenetic hypothesis under consideration, *Cladia applies* only to one of the members of this natural kind in each phylogenetic hypothesis.

ARTICLE 6: Historicism and essentialism in phylogenetic biology, **Y. Bertrand** & M. Härlin, manuscript.

The last article of the compendium starts the discussion where the previous paper finished off and asks how the previous conclusion, namely that phylonames are fixed to natural kinds, can still hold in cases of definitional emendations. Redefinitions pose a threat to any form of natural kind theory, because it implies that past workers were mistaken in supposing that there is such a kind. One of the solutions is a hybrid account between the causal and the descriptive theories of reference (Sterelny, 1983). Such theory supposes that the researcher actually rely on the referent plus some description reflecting the true essence of the kind, when they determine the meaning of names. When the reference is too out of sync with the associated description, the researcher might decide to reground the term into a different entity.

LaPorte (2004) proposes a different solution. He answers that the causal theory is correct, but the claim that we discover the true essence of kinds is wrong. He illustrates with the guinea-pig example which according to some phylogenetic hypotheses is not

part of the clade including all traditionally recognized myomorph rodents. Instead, the guinea-pig and the myomorph rodents diverged before the separation between myomorph rodents and the lineage leading to primates and artiodactyls (Graur *et al.*, 1991). As a consequence, scientists might select any of the following option: a) retain the extension of rodent and accept the conclusion that the guinea-pig is not a rodent, b) adjust the extension and encompass horses and humans into the rodent clade in order to preserve guinea-pig's "rodenthood". LaPorte asserts that any of the above options is perfectly valid, which entails that there is no fact of the matter about which is the right one. This leads him to the anti-Kripkean conclusion that new evidences do not tell us something about what we were referring to before we collected the evidence. From this account, it is clear that past and current workers mean something different when using 'Rodent'. Because often there are several equally good options, we cannot say that in the past it was false that a guinea-pig was a rodent. For LaPorte it was simply vague and we couldn't cogently decide. Now with the change of meaning, it has become precise and according to the selected option we can answer about the truthiness of the statement "the guinea-pig is a rodent".

In our article, we take elements from both the hybrid theory and LaPorte's view. However we challenge LaPorte's argument that the different alternatives facing the taxonomist, were equally good options to select from. We put forth the idea that scientists are driven when making reference change by the will to narrate the same sort of history, in other words, by the will to keep the role of taxa as similar as possible in their successive applications. It doesn't imply that they never take the responsibility of deeply modifying extensions of names. A famous example is *Dinosauria* which used to be the

central subject in some paleontological explanation. Its current extension encompasses extent birds; therefore the name plays a new role in explanations pertaining to some living flying vertebrates.

When the taxonomists' needs are taken into consideration we are forced to modify the ontological picture of taxa presented in the fifth article. A taxon is not a single natural kind defined by one phylogenetic definition, but a collection of several natural kinds corresponding to the successive redefinitions. All together these natural kinds form an entity which we call narrative. In a narrative, integration of the different kinds into a whole proceeds from the systematist's desire to tell a good story.

CONCLUSION

I remember using in my very first article on the PhyloCode (Bertrand & Pleijel, 2003) the worn out discussion between Alice and Humpty Dumpty:

""When I use a word", Humpty Dumpty said in a rather scornful tone, "it means just what I choose it to mean – neither more nor less".

"The question is," said Alice, "whether you can make words mean so many different things." (Carroll, 1872: 124)."

This article concludes by agreeing with Humpty Dumpty that language is a convention but responds to Alice's worry that once a definition is given to a word, their association should remain fixed. Then, I ended by advertising how well the fixity of the association is secured in the PhyloCode, which allows us to cease arguing about definitions and to finally focus on phylogenies.

Nothing in this quotation strikes me as exactly wrong even now; but what strikes me is that I was walking on the razor's edge without knowing it. In the intervening years I have come to see that one cannot come to grips with the real problems in nomenclature and systematics in general without being more sensitive to the philosophical standpoint than I was willing to be when I wrote those words. Becoming more sensitive to that position had consequences which I did not expect. During my thesis, I have encountered that names can refer to individuals, to natural kinds, or even to narratives. I have explored how our knowledge of the world is partly imposed by nature, partly built by consensus between scientists. I have understood that taxa have or lack essential properties depending from where we are observing them. I hope that the reader is now convinced that phylogenetic nomenclature is a fantastic tool when it comes to explore the ontology and epistemology of systematics.

ACKNOWLEDGEMENT

My PhD was fully supported by funds from Södertörns Högskola, which allowed me to work in ideal conditions. I feel really lucky to have been in such stimulating environment.

I would also like to thank some individuals parts of the *Homo sapiens* or *sapiens Homo* (this is not decided yet) species.

First Fredrik Pleijel, without who I wouldn't be writing these words. I would certainly be doing something else, biologist for instance. Thanks for having guided my first steps in science and for trying to teach me some rigor and tenacity. Notice, I said "trying".

Then about Mikael Härlin, what should I say? Where to begin? I could talk about the amazing freedom I had during my work, the wise advices, but it all sounds lame. I've never been talented at expressing gratitude, so I'm just going to quote a dear friend of mine who admitted the other day hoping by the age of 70 to "become as humble and as knowledgeable as Mike". Let's add that my friend has my age, otherwise it doesn't make much sense. For myself, I just hope I won't have to wait that long.

My family for always supporting my choices, my friends, who make such a great difference with this place. I choose my friends because they are different in their views, strange in their ways and because they love me. They are all women and men of some intellectual power, and consequently they all appreciate me. Is that very vain of me? I think it is rather vain.

Without all of you, this world, although possible, wouldn't be conceivable. Thanks.

REFERENCES

Anderson, A.N. 1995. Measuring more of biodiversity: genus richness as a surrogate for species richness in Australian ant faunas. *Biological Conservation* 73: 39-43.

Avise, J.C., & G.C., Johns. 1999. Proposal for a standardized temporal scheme of biological classification for extant species. *Proceedings of the National Academy of Sciences of the United States of America* 96: 7358-7363.

Balmford, A., Green, M.J.B., & M.G., Murray. 1996a. Using higher-taxon richness as a surrogate for species richness: I. Regional tests. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263: 1267-1274.

Balmford, A., Jayasuriya, A.H.M., & M.J.B., Green. 1996b. Using higher-taxon richness as a surrogate for species richness: II. Local applications. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263: 1571-1575.

Balmford, A., Lyon, A.J.E., & R.M., Lang. 2000. Testing the higher-taxon approach to conservation planning in a megadiverse group: The macrofungi. *Biological Conservation* 93: 209-217.

Benton, M.J. 2000. Stems, nodes, crown clades, and rank-free lists: Is Linnaeus dead? *Biological Reviews* 75: 633-648.

Bernier, R. 1984. The species as an individual: Facing essentialism. *Systematic Zoology* 33: 460-469.

Bertrand, Y. & F., Pleijel. 2003. Nomenclature phylogénétique: Une réponse. *Bulletin de la Société Française de Systématique* 29: 25-28.

Brogaard, B. 2004. Species as individuals. *Biology and Philosophy* 19: 223-242.

Boyd, R. 1999. Homeostasis, species and higher taxa. In: Wilson, R.A. (ed.), *Species:*

New Interdisciplinary Essays. Cambridge, Massachusetts: The MIT Press, pp. 141-185.

Cantino, P.D., & K., de Queiroz, 2007. International code of phylogenetic nomenclature, Version 4b. Available at http://www.ohiou.edu/phylocode.

Cantino, P.D., & R.W., Sanders. 1986. Subfamilial classification of Labiatae. *Systematic Botany* 11: 163-185.

Cantino, P.D., & R.G., Olmstead. 2008. Application of phylogenetically defined names does not require that every specifier be present on a tree. *Systematic Biology* 57: 157-160.

Caplan, A. 1981. Back to class: A note on the ontology of species. *Philosophy of Science* 48: 130-140.

Carroll, L. 1872. *Through the looking-glass, and what Alice found there*. New York: MacMillan & Co.

Carpenter, J.M. 2003. Critique of pure folly. *The Botanical Review* 69: 79-92.

Darwin, C. 1859. *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. London: J. Murray.

Dayrat, B., Cantino, P.D., Clarke, J.A., & K., de Queiroz. 2008. Species names in the *PhyloCode* : The approach adopted by the International Society for Phylogenetic Nomenclature. *Systematic Biology* 57: 507-514-

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard, D.J., & S.H., Berlocher (eds.), *Endless Forms: Species and Speciation*. Oxford: Oxford University Press, pp. 57-75.

de Queiroz, K. 1999. The general lineage concept of species and the defining properties of the species category. In: Wilson, R.A. (ed.), *Species: New Interdisciplinary Essays*. Cambridge, Massachusetts: The MIT Press, pp. 49-89.

de Queiroz, K. 2005. A unified species concept and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences* 56: 196-215

de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56: 879-886.

de Queiroz, K., & J., Gauthier. 1990. Phylogeny as a central principle in taxonomy: Phylogenetic definitions of taxon names. *Systematic Zoology* 39: 307-322.

de Queiroz, K., & J., Gauthier. 1992. Phylogenetic taxonomy. *Annual Review in Ecology and Systematics* 23: 449-480.

de Queiroz, K., & J., Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. *Trends in Research in Ecology and Evolution* 9: 27–31.

Dobzhansky, T. 1964. Biology, Molecular and Organismic, *American Zoologist*, 4: 443-452.

Dyke, G.J. 2002. Should paleontologists use “phylogenetic” nomenclature? *Journal of Paleontology* 76: 793-796.

Ereshefsky, M. 2001. *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. New York: Cambridge University Press.

Feyerabend, P.K. 1975. *Against method: Outline of an anarchistic theory of knowledge*. Atlantic Highlands, NJ: Humanities Press.

Gaston, K.I., & P.H., Williams. 1993. Mapping the world's species-the higher taxon approach. *Biodiversity Letters* 1: 2-8.

Ghiselin, M.T. 1966. On psychologism in the logic of taxonomic controversies. *Systematic Zoology* 15: 207-215.

Ghiselin, M.T. 1969. *The triumph of the Darwinian Method*. Berkeley: University of California Press.

Ghiselin, M.T. 1974. A radical solution to the species problem. *Systematic Zoology* 23: 536-544.

Ghiselin, M.T. 1997. *Metaphysics and the origin of species*. Albany, NY: State University of New York Press,

Goodman, N. 1973. Fact, Fiction, and Forecast. 3rd ed. Indianapolis & New York: Bobbs-Merrill.

Graur, D., Hide, W.A., & W.H., Li. Is the guinea-pig a rodent? *Nature* 351: 649-652.

Griffiths, G.C.D. 1976. The future of Linnaean nomenclature. *Systematic Zoology* 25: 168-173.

Griffiths, P.E. 1994. Cladistic classification and functional explanation. *Philosophy of Science* 61: 206-227.

Hennig, W. 1966. *Phylogenetic systematics*. Urbana: University of Illinois Press.

Hibbett, D.S., & M.J., Donoghue. 1998. Integrating phylogenetic analysis and classification in fungi. *Mycologia* 90: 347-356.

Hickman, P.H.Jr., Roberts, L.S., Larson, A., l'Anson, H., & D.J., Eisenhour. 2006. *Integrated Principles of Zoology*, 13th ed. New York: McGraw-Hill.

Hilpinen, R. 1976. Approximate Truth and Truthlikeness, In: Przel/ecki, M., Szaniawski, K., & R. Wójcicki (eds.), *Formal Methods in the Methodology of Empirical Sciences*, D. Reidel, Dordrecht: Ossolineum, Wrocl/aw, pp. 19-42.

Hull, D. 1976. Are species really individuals? *Systematic Zoology* 25:174-191.

Härlin, M. 1998. Taxonomic names and phylogenetic trees. *Zoologica Scripta* 27: 381-390.

Kaufman, M., & M., Wink. 1994. Molecular systematics of the Nepetoideae (family Labiatae): phylogenetic implications from rbcL gene sequences. *Zeitschrift für Naturforschung* 49c: 635-645.

Kitcher, P. 1984. Species. *Philosophy of Science* 51: 308-333.

Kitts, D.B., & D.J., Kitts. 1979. Biological species as natural kinds. *Philosophy of Science* 46: 613-622.

Kojima, J. 2003. Apomorphy-based definition also pinpoints a node, and PhyloCode names prevent effective communication. *The Botanical Review* 69: 44-58.

Kripke, S.A. 1963.Semantical analysis of modal logic I, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9: 67–96.

Kripke, S.A. 1980. *Naming and Necessity*. Harvard: Harvard University Press.

Kuhn, T.S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

LaPorte, J. 2004. *Natural kinds and conceptual change*. New York: Cambridge University Press.

Lee, M.S.Y. 1997. Documenting present and past biodiversity: conservation biology meets palaeontology. *Trends in Ecology and Evolution* 12: 132-133.

Lee, M.S.Y, & M., Wolsan. 2002. Integration, individuality and species concepts. *Biology and Philosophy* 17: 651-660.

Lewis, D.K. 1968. Counterpart theory and quantified modal logic. *Journal of Philosophy* 65: 113-126.

Lewis, D.K. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Mishler, B.D., & R.N., Brandon. 1987. Individuality, pluralism, and the phylogenetic species concept. *Biology and Philosophy* 2: 397-414.

Moore, G. 2003. Should taxon names be explicitly defined? *The Botanical Review* 69: 2-21.

Nixon, K.C., & J.M., Carpenter. 2000. On the other "Phylogenetic Systematics" *Cladistics* 16: 298-318.

Nixon, K.C., Carpenter, J.M., & D.W., Stevenson. 2003. The PhyloCode is fatally flawed, and the "Linnaean" system can easily be fixed. *The Botanical Review* 69: 111-120.

Putnam, H. 1975 *Mind Language and Reality; Philosophical Papers*, volume 2, Cambridge: Cambridge University Press.

Rieppel, O. 2005. Monophyly, paraphyly, and natural kinds. *Biology and Philosophy* **20**: 465-487.

Salmon, N. 2005. *Reference and Essence*. 2nd ed. Amherst, New York: Prometheus Books.

Sibley, C.G., & J.E., Ahlquist. 1990. *Phylogeny and classifications of birds*. Yale University Press, New Haven-London.

Sober, E. 1980. Evolution, Population Thinking, and Essentialism. *Philosophy of Science* 47: 350-383

Sober, E. 1993. Evolution, population thinking and evolution. In: Sober, E, (ed), *Conceptual issues in evolutionary biology*. 2nd ed. Cambridge, Massachusetts: The MIT Press, pp. 161-190.

Sober, E. 2000. *Philosophie of Biology*. Boulder& Oxford: Westview Press.

Sterelny, K. 1983. Natural kind terms. *Pacific Philosophical Quarterly* 64: 110-125.

Stevens, P.F. 1994. *The development of biological systematics*. New York: Columbia University Press.

Warwick, R.M. 1988. The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. *Marine Pollution Bulletin* 19: 259-268.

Williams, P.H., & K.J., Gaston. 1994. Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness? *Biological Conservation* 67: 211-217.